Predicting Ecosystem Pattern across Landscapes

Sean Basquill[1] and Shawn Leroux[1]

[1]Memorial University of Newfoundland

November 15, 2022

Abstract

Predictive modelling is fundamental to ecology and essential for objective biodiversity assessment. However, while predictive biodiversity models are generally well-developed, models for predicting patterns within and among ecosystems have not been adequately operationalized. We contend the scarcity of such models marks a concerning gap in the scientific community's ability to make ecosystem predictions across landscapes, and more broadly for supporting the conservation of biodiversity and ecosystem functions. We propose ecosystem spatial pattern models (ESPM) to fill this gap in modelling capacity. Under our approach to ESPM, spatial patterns of ecosystem properties are the basis for resolving ecosystem organization at local and landscape extents. Our integrative modelling framework differs from others in that it accords biotic and abiotic constituents equally, based on with their joint mechanistic influence on ecosystem dynamics. Development of ESPM is especially timely for ecosystem assessment is undergoing a contemporary groundswell, as scientists and conservation groups propose ambitious targets for ecosystem conservation and restoration.

Abstract

Predictive modelling is fundamental to ecology and essential for objective biodiversity assessment. However, while predictive biodiversity models are generally well-developed, models for predicting patterns within and among ecosystems have not been adequately operationalized. We contend the scarcity of such models marks a concerning gap in the scientific community's ability to make ecosystem predictions across landscapes, and more broadly for supporting the conservation of biodiversity and ecosystem functions. We propose ecosystem spatial pattern models (ESPM) to fill this gap in modelling capacity. Under our approach to ESPM, spatial patterns of ecosystem properties are the basis for resolving ecosystem organization at local and landscape extents. Our integrative modelling framework differs from others in that it accords biotic and abiotic constituents equally, based on with their joint mechanistic influence on ecosystem dynamics. Development of ESPM is especially timely for ecosystem assessment is undergoing a contemporary groundswell, as scientists and conservation groups propose ambitious targets for ecosystem conservation and restoration.

Introduction

Predictive modelling is essential for understanding the natural world and for making informed decisions about current and future states. Dramatic biodiversity shifts, compounded with the poorly understood ecological consequences of emerging climatic regimes, have served as a clarion call for more predictive conservation science. These shifts have also helped to reaffirm the broader role of prediction in basic and applied ecology (Houlahan et al. 2017). However, while predictive biodiversity models (e.g., species or community distribution models) are generally well developed, and frequently employed to inform conservation decision-making, these advancements have not been fully realized in predictive spatial models of ecosystems (Fulton et al 2019, Geary et al 2020, Halvorsen et al 2020, Simensen et al 2020). Spatially explicit models are required for predicting intra- (e.g., structure, composition, function, spatial properties) and inter-ecosystem (e.g., identity, spatial properties) patterns across landscapes. The scarcity of such models hinders the scientific community's ability to make objective ecosystem predictions and near-term forecasts at sub-global scales.

Predictive spatial models have emerged as the principal tool to understand biodiversity patterns, and to forecast the effects of ecological change (Araújo et al 2019). While the scope of these models has progressively broadened to address multiple components of biodiversity, biotic predictions are by far the most common application (Pollock et al 2020). In their review of spatially explicit biodiversity models, Zurell et al (2022) report 80% of published models are focused on species distributions. A smaller percentage of predictive spatial models address higher order (e.g., communities) components of biodiversity but these too are, almost exclusively, biotic entities (Pollock et al 2020, Zurell et al 2022). For example, joint distribution and generalized dissimilarity models of communities (e.g., community distribution, diversity, or interactions) are common extensions of predictive species distribution models (Zurell et al 2020, Wilkinson et al 2021, Mokany et al 2022). In sharp contrast, comparable spatial models of ecosystem pattern – and more specifically those models that integrate biotic and abiotic variables and system-level properties (Table S1) – are extremely rare (Dor-Haim et al 2019, Halvorsen et al 2020, Simensen et al 2020). This gap in modelling capacity is remarkable. To fill this gap, new spatial modelling approaches are needed for modelling ecosystem patterns across landscapes. Development of suitable models is especially timely for ecosystem assessment (e.g., Keith et al 2022) has undergone a contemporary groundswell, as the scientific community (e.g., Watson et al 2020, Bullock et al 2022) propose ambitious targets for ecosystem conservation and restoration.

Conservation practitioners and regulatory authorities rely on spatial models to ensure decisions are informed by robust predictions and related measures of statistical uncertainty. Most conservation decisions are realized in geographic space and selecting the appropriate spatial scale to maximize success is imperative (Di Minin et al 2022). Trend analyses show the ecosystem concept is frequently employed to structure spatially based conservation plans and to mitigate stressors including climate change (Anderson et al 2021). The most prominent commitments to ecosystem conservation have been established in global agreements (e.g., Convention on Biological Diversity 2021), while implementation falls to national signatories and to sub-national governments and conservation organizations (Perino et al 2022). In Canada, for example, standardized spatial data of significant 'key biodiversity areas' including threatened ecosystems, have been recommended to inform national conservation planning (Shea et al 2018). However, achieving this ecosystem objective presents a challenge in Canada that is shared elsewhere. The challenge is that spatially explicit ecosystem prediction is recognized as an essential tool to inform regional (i.e., continental, national, sub-national) commitments to global conservation initiatives but very few fully integrated spatial models of ecosystem pattern, at regional landscape scales, are available. Some members of the scientific community have raised this as a problematic issue (Fulton et al 2019, Geary et al 2020), but it is seldom acknowledged, particularly in the applied literature (Halvorsen et al 2020).

In this Forum, we propose a novel framework for spatially explicit modelling of ecosystem pattern. Our framework differs from other spatial ecosystem models in that it accords biotic and abiotic components equally, commensurate with their joint influence on ecosystem assembly (sensu Keith et al 2022). Under this approach, the ecosystem is modelled as a spatially explicit place demarcated by the strength of concordance among constituent biotic and abiotic variables (Box 1). Different ecosystems are in turn discriminated by relatively unique combinations of biotic and abiotic variables that recur along spatial gradients (Figure 1E, 1F). Under our approach, groups of local ecosystems scale up to form the heterogenous landscapes within which they occur (Figure 1F). The bounds of an ecosystem's local extent can only be determined by considering both the extent and identity of proximate ecosystems. Consequently, the spatial organization of ecosystems, comprising a particular landscape, can be predicted by identifying spatially structured differences among constituent ecosystem types. Our framework offers a quantitative strategy for modelling the range-wide occurrence, identity, and characteristic properties of individual ecosystems. Furthermore, it provides an approach for determining regional patterns of ecosystem spatial organization, and related changes in biotic-abiotic relationships, across heterogenous landscapes.

This Forum includes four primary components. First, we briefly review existing ecosystem modelling approaches, with a focus on models formulated to predict ecosystem or meta-ecosystem spatial patterns at regional landscape extents – the spatial extent where most terrestrial environmental decisions are made. Second, we demonstrate how a limitation in predictive modeling capacity occurs because of disparate scientific traditions and related conventions of modelling practice. Third, we outline a novel conceptual framework for predictive ecosystem spatial pattern modelling at regional landscape extents. Our framework includes an analytical workflow illustrated with a case study. Lastly, we conclude with a perspective where we propose future research directions and opportunities.

Spatial Ecosystem Models – a Brief Overview

Most ecosystem models lie on a continuum of basic to applied objectives with related modelling trade offs (DeAngelis and Yurek 2016). Basic ecosystem models are often mathematical models, primarily intended to develop, test, and formalize ecosystem theory (Jørgensen 2016). Most of these models have been aspatial, spatially implicit, or limited to few geographic sites and ecosystems (DeAngelis and Yurek 2016). The other end of the spectrum is marked by applied ecosystem models, which are usually fit to empirical data (Evans 2012) or derived using expert-based protocols (Halvorsen et al 2020). Applied models are normally intended for description or prediction of ecosystem conditions, including their specific arrangement in geographic space (Pickett and Candenasso 2002, Evans 2012, Geary et al 2020). Recently though, increasing awareness of the importance of space in shaping ecosystem pattern and process has prompted appeals to extend spatially explicit ecosystem modelling more evenly across scientific domains (e.g., DeAngelis and Yurek 2016, Fulton et al 2019). Spatially explicit approaches are not necessary in all ecosystem models, but they are indispensable for models intended for making predictions of ecosystems across landscapes (Evans 2012). We highlight the importance of spatially explicit modelling approaches as a strategy to help bridge theoretical and applied branches of ecosystem science. Furthermore, we demonstrate their use for contemporary and near-future ecosystem predictions, at local to regional landscape extents.

Ecosystems occur in geographic space (Pickett and Candenasso 2002) and space has long been recognized as a fundamental property mediating the expression of ecosystem heterogeneity (Rowe 1961, Holling 1992). Yet until recently, advances in spatially explicit ecosystem modelling have occurred under the somewhat independent auspices of ecosystem and landscape ecology (Loreau et al 2003). The latter discipline has commonly considered ecosystems and ecosystem processes as landscape components, focusing on their contributions to spatial heterogeneity (Pickett and Cadenasso 1995). Ecosystem ecologists have, in turn, mainly concentrated their attention on the exchange of organisms, energy, materials – particularly resources such as inorganic nutrients – along continuous geographic gradients or discretized landscape mosaics (Massol et al 2011, Gounand et al 2018). Their burgeoning approach – sometimes called '*spatial or landscape ecosystem ecology* ' (sensu Loreau et al 2003) – often uses the meta-ecosystem as an analytical construct (Massol et al 2011) and frequently relies on mathematical models (e.g., Harvey et al 2021). However, current perspectives in spatial ecosystem ecology (e.g., Leroux et al 2017, Gounand et al 2018, Harvey et al 2021) call for improved integration of ecosystem and meta-ecosystem models with real-world landscape contexts to inform conservation policy. Thus far, this synthesis of landscape and ecosystem ecology has predominantly been championed through spatially explicit models of ecosystem functional properties (e.g., Leroux and Loreau 2012, Gounand et al 2018, Soranno et al 2019). We contend further synthesis is needed to predict other spatially structured aspects of ecosystem variation, including common pools of biotic and abiotic complexity at local and regional landscape extents. Furthermore, those biotic and abiotic factors interacting in a particular time and place, influence location-specific patterns of ecosystem compositional, structural, and functional properties (Nash et al 2014). The spatially explicit ecosystem modelling strategy presented in this forum is structured to illustrate these interdependent aspects of intra- and inter-ecosystem organization and variation.

Scientific Traditions in Ecosystem Ecology

Ecosystems are complex and may include hundreds or even thousands of biotic and abiotic components (Loreau 2010). Studying or modelling all aspects of ecosystem complexity is not possible or desirable. To represent ecosystem variation, one needs a basis for simplification and this simplification comes with trade-off (Geary et al 2020). In the basic ecosystem sciences, the customary emphasis on ecological function (Box 2), has come at the expense of detailed information on biotic properties (Loreau 2010, Mokany et al 2016) and on the contributing roles of abiotic complexity (Johnson and Martin 2016, Richter and Billings 2015, Dor-Haim et al 2019, Hjort et al 2022). Many of these traditions are long-standing and mirrored across general ecology (Loreau 2010). Loreau (2010) attributes them to a divergence of practice between community and ecosystem ecologists. Species interactions and patterns are largely studied by community ecologists. Ecosystem ecologists instead emphasize interactions between species and their abiotic environments, particularly those interactions that give rise to fluxes of materials. These latter functions (Box 1) are widely considered the fundamental foci for basic modelling and empirical inquiry at the ecosystem level. While biota and biotic processes are considered ecosystem components, they are normally relegated to the community (Loreau 2010) and generally interpreted as components of biodiversity. These enduring traditions of practice are also echoed across the spectrum of basic and applied ecosystem ecology, and their recognition within this broader scope of ecosystem science is not new (e.g., Rowe 1961, Blew 1996).

Blew (1996) outlined three major traditions of ecosystem science each rooted in relatively distinct perceptions of the ecosystem concept and its salient features (Box 2). The three traditions place differing levels of emphasis on biotic, abiotic, or functional ecosystem features (Table S1). While Blew's (1996) summary has not been updated to include modern theory and modelling approaches, some of which strive for improved integration (e.g., Geary et al 2020), similar conventions in ecosystem science exist today (Baveye et al 2018, Zarnetske et al 2019, Alahuhta et al 2020). Blew (1996) attributes such divergence of scientific practice to controversies surrounding the ecosystem concept itself, which Loreau (2020) shows have pervaded since its inception. Many of the debates originate because the ecosystem concept is at once both comprehensive and flexible (Pickett and Candenasso 2002, Currie 2011). The ecosystem concept's inherent flexibility lends itself to different approaches for modelling wholeness and integration among key unit components (Pickett and Candenasso 2002, Gignoux et al. 2011, Jørgensen 2016). However, part of the debate with past modelling efforts have been discussions over which system-level properties (Table S1) best exemplify wholeness, adequately represent complexity, and underpin quantifications of variation (Currie 2011). We concur with Jørgensen (2016) and suggest there are many departure points for simplifying the ecosystem, as an object for modelling, to help address these considerations, and to advance understanding of ecosystem patterns. Furthermore, we highlight the importance of spatially explicit approaches as a distinct strategy for bridging different analytical traditions, and for producing predictive models needed for ecosystem forecasts. These ecosystem modelling trade-offs will necessarily vary according to purpose. And while research to distinguish ecosystem modelling strategies and to assess their relative merits are on-going (e.g., Geary et al 2020), our novel approach falls outside the scope of recent efforts and related recommendations.

In the next section, we present a framework for ecosystem conceptualization and spatial pattern modelling. The framework integrates elements of the ecosystem science traditions outlined above and presented in Box 2. It does not emphasize biotic, abiotic, or functional features of ecosystem variation. Instead, our framework focuses on the causes and consequences of these features' concordance (Box 1) in space and time. Additionally, we show how that spatial concordance shapes patterns of ecosystem heterogeneity at local and regional landscape extents.

Integrative Ecosystem Spatial Pattern Modelling

Our strategy for ecosystem spatial pattern modelling is summarized as a conceptual framework. We draw concepts from ecosystem, community, and landscape ecology, seeking to demonstrate an integrative basis for predicting the drivers and outcomes of ecosystem assembly. To illustrate practical application of our framework, we provide a workflow (Figure 1) based on a case study from Cape Breton, Nova Scotia, Canada. We showcase data, analytical procedures, and interpretive guidelines required for framework operationalization, as they relate to the case study. These operational considerations are applicable to any predictive spatial model of regional ecosystem pattern. In that regard, an overarching objective of our framework is to help catalyze new perspectives and statistical methods for obtaining general insight into the origin and implications of ecosystem spatial organization. We begin by summarizing the conceptual foundation of our framework.

Conceptual foundation

The premise of our modelling strategy is that ecosystems can be predicted by resolving patterns of relative spatial concordance (Box 1) among biotic and abiotic features (Table S1). The particular combination of features found along a geographic gradient, and the spatial relationships among them (Figure 2), provide an integrated basis for distinguishing conterminous ecosystems and their respective boundaries (Figure 2). Such distinguishing properties are not absolute, for ecosystems intergrade and vary continuously in space and time (Leroux and Loreau 2012, Newman et al 2019, Soranno et al 2019). These gradations occur because individual biotic-abiotic features, intrinsic to a particular ecosystem, may transcend boundaries of locally adjacent ecosystems. However, ecosystem-level organizational patterns can be identified by relative differences in the collective strength of concordance among constituent features. For example, in boreal landscapes, tree species composition may be relatively concordant across adjacent upland forest and treed bog ecosystems. Here, key differences are expressed in lower vegetation layers, and in soil, organic matter, hydrology, and topographic properties. Under this approach, different ecosystems can be discriminated by relatively unique and consistent combinations of features, that recur predictably across landscapes. Landscapes of heterogeneous ecosystems are composed of these different ecosystems. Here, each ecosystem manifests at characteristic spatial scales (Massol et al 2011, Soranno et al 2019, Bullock et al 2022) and has characteristic levels of integration (Rowe 1961, Loreau 2020), uniformity (Pickett and Candenasso 1995, Halvorsen et al 2021), and distinction (Dor-haim et al 2019, Simensen et al 2020).

A nuance of our interpretation is that we do not seek a means to reconcile patterns of low spatial concordance between or among biotic and abiotic ecosystem constituents. Instead, we emphasize levels of collective spatial concordance within and across biotic-abiotic domains (Box 1). Furthermore, we suggest collective levels of concordance are a relative, continuous, and changeable ecosystem property. Some ecosystems have inherently higher concordance, whereas others are characterized by lower relative levels of this emergent ecosystem property (Loreau 2020). Holling (1992) attributed differences in overall concordance to keystone variables. Keystone variables are individual ecological processes or constituents with a disproportionate influence on ecosystem spatial organization and variation (Holling 1992). Over time, these variables, many of which operate across scales, can catalyze developmental changes in other aspects of ecosystem structure, composition, and function (Nash 2014, Newman et al 2019). Such changes often lead to more consistent expressions of landscape pattern and spatial grain (Peterson 2002), and potentially to higher levels of ecosystem integration (Rowe 1961, Loreau 2020).

Case Study

Overview

We define ecosystem spatial pattern models (ESPM) as statistical models formulated to predict spatially structured dimensions of intra- and inter-ecosystem variation. We adapt elements of community-level spatial models (see review in D'Amen et al 2017) to configure our analytical strategy for ESPM. Community-level models are multi-species extensions of species distribution models (Norberg et al 2019, Zurell et al 2020), formulated to simultaneously model the presence or abundance of species comprising communities, or to model patterns of species compositional dissimilarity among sites, at various scales of resolution (Nieto-Lugilde et al 2018). These models are also commonly applied to predict community-level properties (e.g., species richness and composition – Norberg et al 2019) and or assemblages (e.g., community type – Pinto-Ledezma and Cavender-Bares 2021) across geographic gradients. Ecosystem spatial pattern models differ from community-level distribution models in three main respects: modeled response variables, prerequisite test and predictor data, and the specific assembly determinants underpinning model prediction. We briefly summarize these features below in the workflow applied in our case study.

To build our strategy for ESPM, we focus on one of the three main strategies to community-level modelling (see Ferrier and Guisan 2006). This approach, referred to as *assemble and predict together* , models species-environmental relationships concurrently with spatial predictions of community-level properties and or entities (Ferrier and Guisan 2006, D'Amen et al 2017). Our analytical approach most closely mirrors joint-species distribution modelling, a strategy to predict the responses of multiple species occurring throughout a study space (see review in Warton et al 2015). We extend this to model ecosystem spatial patterns.

Under our approach to ESPM, spatially concordant patterns of biotic-abiotic co-occurrence and or abundance are the basis for identifying ecosystems at local and landscape extents. To illustrate implementation of an ESPM, we outline a typical modelling workflow (Figure 1) and provide an overview of key procedures. Key components include aspects of ecosystem survey (Figure 1A); ESPM training and predictor data compilation and refinement (Figure 1B), and finalization (Figure 1C); model building (Figure 1D); spatial prediction of individual ecosystem constituents and their properties (Figure 1E); and lastly, prediction of ecosystem-level patterns and properties at local (i.e., within individual ecosystems) and landscape (i.e., among ecosystems or ecosystem types)(Figure 1F) extents. Below we summarize these ESPM components as they relate to our case study. We also provide a detailed overview of the more prominent challenges (see Table 1 for complete list) encountered during model implementation, offering practical solutions to address these obstacles. Our overall intent is not to report on ESPM outcomes, as they relate to our case study, but to illustrate key aspects of model set-up, operationalization, and problem resolution.

Ecosystem Survey

Field data employed in our case study (Figure 1A) were obtained from ecosystem surveys of 1004 plots across Cape Breton (10,311 square kilometers), Nova Scotia, Canada (see general overview of ecosystem survey methods in section S2). Ecosystem survey protocols incorporate different methods for biotic and abiotic variables. These methods may also be applied across multiple spatial extents and grains. Survey methods have a strong influence on the properties and structure of resulting data, with implications for how these data are used in modelling spatial ecosystem patterns. These implications are discussed below in our overview of ESPM training and predictor data.

Data Compilation and Refinement

Training Data

Data required for ESPM include training and predictor datasets. Training (sometimes called test or response) data (Figure 1B) are derived from ecosystem field surveys, as described above. These data must undergo similar preparations, refinements, and partitioning recommended for species (see Guisan et al 2017) and community-level modelling (see Ovaskainen and Abrego 2020, Mokany et al 2022), as applicable. A notable distinction is that ecosystem training data include both biotic and abiotic variables. Differences among these respective survey variables introduce incompatible data structures in the training pool, with a bearing on predictive model outputs. We provide a brief overview of these issues and suggest options to address them. Last, we introduce measures to finalize training data selected for model fitting.

Different measurement scales and units are typically employed for quantifying biotic and abiotic phenomena in field surveys. In our study, vascular plant, bryophyte, and lichen taxa were measured by estimating areal cover by individual species. These estimates were expressed as a percent of total cover for individual vertical vegetation strata (e.g., canopy tree, woody shrub, tree sapling and seedling, herbaceous, lichen, and bryophyte). In contrast, abiotic attributes were recorded using a variety of measurement scales and units, depending on the attribute. These abiotic records include measures in linear (e.g., humus depth in centimeters), categorical (e.g., landform type), concentration (e.g., milligrams/kilogram of calcium in soil), volume (e.g., above-ground woody biomass in cubic meters), ratio (e.g., percent slope), or logarithmic (e.g., pH) units. Pooling ecological variables with disparate measurement scales, units, and limits misrepresents their statistical relationships to one another, and introduces interoperable dimensions into a training dataset (König et al 2019). One possible means to alleviate this problem is to convert all variables to presence/absence records. This method was particularly effective for the types of categorical variables we have in our data (e.g., species identity and abiotic attributes such as humus, soil, and bedrock type) but less appropriate for continuous measures, particularly those which were present at all surveys locations (e.g., vegetation height). To address this latter problem, ubiquitous variables can be eliminated from the data pool, a procedure implemented on biotic training data employed in community-level modelling (e.g., Ovaskainen and Abrego

2020). Alternatively, some continuous data (e.g., pH) could be converted to categorical classes, employing empirically supported class intervals. Given the diverse data types in our case study, we excluded ubiquitous variables and fit a hurdle model (see Ovaskainen and Abrego 2020) with and without abundance data.

A second challenge arising in training data compilation is the disproportionate ratio of biotic to abiotic variables recorded in field surveys. Tens or hundred of species are frequently documented from individual ecosystem plots, whereas the number of abiotic attributes assessed (Figure 1A) is rarely comparable. Our case study plot data are comprised of abundance records of over 900 vascular plant, lichen, and bryophyte species, coupled with only 20 abiotic variables recorded at the site-level and 10 to 30 additional abiotic measurements taken from soil profiles at each survey location (see section S2). Imbalanced numbers of biotic and abiotic data survey variables unduly weights biota in training data pools, with implications for the relative influence these variables have on predictions of joint concordance among biotic and abiotic ecosystem constituents. One option to address this imbalance is to aggregate species into groups based on common ecological traits (i.e., life history, morphological, or physiological characteristic of an individual) and to employ the resulting trait groups to represent biota in a test data pool. Incorporating species traits into predictive spatial biodiversity models has been shown to greatly improve model performance (Regos et al 2019). Traits can also be invaluable for understanding species response to their environment and their ecological roles in communities and ecosystems (Kissling et al 2018). One solution we implemented in our case study was to aggregate species with common traits to the genus level; an example was *Sphagnum* , a species-rich genus commonly found on moister soils.

Methods outlined above can be applied with different subsets of training data. This approach can help reveal ecosystems at different scales, and to determine whether separate models, with distinct pools of training data, are required for predicting sub-components of total ecosystem diversity in a study landscape. Variations in training data structure and origin can strongly shape spatial biodiversity model predictive performance (Guillera-Arroita et al 2015, Mod et al 2020). In our case study, pools of test data were scaled across levels of ecological complexity to determine how variations in training data dimensionality affected model performance. For example, differences in growth-form (tree, shrub, herb) dominance, surficial origin (glacial, aeolian, marine, lithic), soil properties (upland, wetland soils), and biogeoclimatic regionality (alpine, boreal, temperate) were employed to parse training data into less complex subsets.

Predictor Data

While training data employed in spatial biodiversity modelling are typically obtained from field surveys, predictor (sometimes called environmental, explanatory, or covariate) data are derived from remotely sensed, or modelled, spatial data extending completely across a study region (Bryn et al 2021). These data are required to model ecosystems continuously in space (Figure 1B). Very few studies have evaluated the suitability of predictor data for this type of predictive modelling (Dor-Haim et al 2019, Halvorsen 2020, Simensen 2020). We suggest suitable predictors should encompass spatial data on biotic and abiotic variables, which are independent from survey records (i.e., training data). They may also include spatial data on functional ecosystem properties (e.g., productivity), physiognomic structure (e.g., dominate growth form), and on regular or episodic disturbance (e.g., flooding, forestry, fire). Potential predictor datasets must be refined to address the same limitations (e.g., correlation, collinearity, and variance inflation) presented in modelling species and communities; we do not review these measures (see Guisan et al 2017 for guidance). Here we focus on the factors considered to finalize the selection of predictors (Figure 1C) employed in our case study, and the implications of these factors for modelling ecosystem pattern.

To assemble our pool of candidate predictors, we considered three factors – ecological response, assembly mechanisms, and spatial scale(s) of influence. In general, we sought predictors deemed important for shaping variations in the distribution or prominence of those ecosystem features (Table S1) which scale up to drive spatial ecosystem patterns. This approach is analogous to the basis of joint species distribution modelling, which assumes species respond mutually to both their physical environment, and to one another (D'Amen et al 2017, Ovaskainen and Abrego 2020). We extend this idea to encompass reciprocal influences among biotic and abiotic variables and sought covariates (Table S3) to predict singular and joint biotic-abiotic responses.

Many ecological assembly drivers (e.g., climate, topography, surficial geology) affect both biotic and abiotic ecosystem constituents and facilitate interactions between them. In our study region, a challenge we encountered is that some predictors are only available at relatively coarse-grains (e.g., climate - approximately 1 kilometer resolution) precluding their use in our compilation of finer-grained predictors. Our initial approach was to seek predictors with spatial grains similar to our survey grain (between 25 and 400 m^2, depending on the ecosystem). Ecosystem surveys were structured around ecologically conspicuous breaks in topographically controlled gradients (e.g., local vegetation), so we employed direct (e.g., remotely sensed vegetation indices) and indirect predictors (e.g., terrain) of those gradients to compile our pool of fine-grained (50 m resolution) candidate predictors (Table S3). Predictor data can be subsequently upscaled to coarser spatial resolutions (e.g., 100, 250, 500, 1000 m) to test the influence of varying spatial grains on predictive model performance. Konig et al (2021), for instance, have demonstrated how joint species distribution model performance varies at different spatial grains. Lastly, investigating the effect of scope, defined as the ratio of spatial extent to grain (Frazier 2022), could provide insights into scaling relationships and their role in predictions of ecosystem pattern.

The combination of predictor data employed in a spatial biodiversity model can be finalized (Figure 1C) before or with model fitting. Where model fitting is used to assist with data selection, selection criteria can be conceptual, theoretical, and or statistical (Guisan et al 2017). Several ecologists (e.g., Araujo and Guisan 2006) have suggested that many biodiversity distribution models do not adequately consider relevant theory to guide predictor data selection (e.g., Poggiato et 2021). Araujo and Guisan (2006) further suggest that greater attention needs to be given to the explanatory value of each predictor, and to its relevance as a causal determinant of ecological pattern. We employed an iterative approach based on model performance, and drew on theoretical and conceptual criteria, to finalize predictor variables employed in our case study (Table S3).

Model Building and Assessment

Model building includes technical aspects of model algorithm selection, managing statistical assumptions, model fitting, and translating model outputs to mapped summaries, among related considerations (Guisan et al 2017, Araujo et al 2019, Zurell et al 2020). In this section we provide an overview of key factors relevant for ESPM algorithm selection. Additional issues, pertaining to ESPM building, are outlined in Table 1.

Algorithm Selection

Previous efforts to model ecosystems in space have largely emphasized biotic, abiotic, *or* functional response variables (Box 1, Table S1) to predict ecosystem-level variation (Table S2). Our case study goals differ as we emphasize biotic *and* abiotic variables equally. Furthermore, we aim to “assemble and predict together” (sensu Ferrier and Guisan 2006) (see case study overview). Given these conditions, we sought candidate ESPM algorithms which could model biotic and abiotic ecosystem constituents simultaneously and predict individual and shared responses of each variable. In addition, we fit ESPM with in situ (i.e., recorded in the field) training data, thereby better capturing the range of biotic and abiotic complexity characterizing ecosystems. By applying this integrative strategy for ESPM, ecosystems are treated as an *emergent function* (sensu Nieto-Lugilde et al 2018) of biotic-abiotic co-occurrence and patterns of local concordance.

Our initial challenge was defining criteria to select an algorithm from the numerous examples (Table 2) applied for spatial biodiversity modelling. Following our evaluation, we screened algorithms according to their flexibility (e.g., regarding data inputs), implementation (e.g., ease of application), analytical properties, and performance (Table 3). Overall, we emphasized algorithms with the capacity to accommodate pools of presence/absence, or abundance, records of both biotic and abiotic variables as training data; employ species traits; and to allow for interactions among predictors. We also prioritized algorithms which could predict individual and joint responses in space. To assist with our selection, we drew on recent review articles (Nieto-Lugilde et al 2018, Norberg et al 2019) and individual model algorithm assessments (e.g., Warton et al 2015, Wilkinson et al 2021) to identify spatial algorithms with relatively high predictive power. Candidate algorithms that met our requirements include joint species distribution modelling (implemented with Hierarchical Modelling of Species Communities (HMSC)) (Ovaskainen et al 2017), generalized dissimilarity modelling (Mokany et al 2022) and probabilistic bioregion modelling (Hill et al 2020). Of these three algorithms, we selected HMSC for our case study as recent reviews demonstrate its overall flexibility and moderate to high predictive performance for spatial biodiversity modelling (Warton et al 2015, Zhang et al 2018, Norberg et al 2019). Furthermore, availability of a comprehensive methodological guide with tutorials in R (Ovaskainen and Abrego 2020), training modules, and communication (Ovaskainen, Tikhonov pers comm) facilitated implementation for our purpose.

Prediction of Intrinsic Ecosystem Constituents and Properties

Separating individual (Figure 1E, e.g., biotic or abiotic features – Table S1) and shared (Figure 1F, i.e., two or more ecosystem features) responses provides a means to predict both intra- and inter-ecosystem patterns. This approach also improves understanding of commonalities, potential interactions, and properties (e.g., traits) among those biotic and abiotic variables characterizing specific ecosystems or sites. Joint distribution models, such as Hierarchical Modelling of Species Communities (HMSC), assume that responses to environmental circumstances are shared; as such, these models are structured to predict both individual and joint responses among test variables (Ovaskainen and Abrego 2020). Furthermore, HMSC is specifically designed so models of individual response variables are statistically connected, so that information from lower-order (i.e., individual biota and abiotic features) predictions informs higher-order (i.e., aggregate) predictions.

Ecosystem-level Prediction

The ultimate purpose of ESPM is to make ecosystem predictions at both local and regional extents (Figure 1F). A key assumption of this objective is that disparate local ecosystems can be resolved, and that the arrangement and spatial properties of those individual ecosystems determine regional landscape structure and composition. Under these interrelated premises, the diversity and spatial organization of ecosystems, comprising a particular landscape, can be predicted by determining spatially structured differences among constituent ecosystem types. To achieve this ESPM objective, we propose analytical techniques to identify statistical commonalities among groups of biotic and abiotic variables co-responding in predicted geographic space (Table 1). These techniques provide a basis to identify ecosystem types, which we define as ecosystems that recur predictably across landscapes. Technical conventions and methods employed for aspatial classification of ecological units (e.g., community, habitat, landscape types) are suitable for this purpose. A number of these methods, and other less common techniques, have been applied to variables modelled in space (see Table 1).

An issue we considered was how to define statistically defensible thresholds of similarity or uniqueness among ecosystems. Ecosystems are spatially continuous, so the issue centers on quantifying patterns of relative discontinuity, arising from discordant patterns of abundance and or occurrence among ecosystem features jointly modelled in space. Outputs of spatial models are often continuous (e.g., probability of occurrence or co-occurrence) and many modelling applications require the identification of suitable thresholds for identifying ecological meaningful levels of relative continuity or discontinuity across space (Zurell et al 2020). For example, users may want to know if the probability of occurrence, for a particular species, is high enough to justify conservation actions (Bryn et al 2021). Thresholds have also been applied for determining whether co-occurrence patterns among groups of species reveal higher order levels of spatially structured biodiversity (e.g., communities - Ferrier and Guisan 2006) or ecological geography (e.g., bioregions - Hill et al 2020). For ESPM, and other spatially structured models of ecological diversity and organization, the challenge can be framed as discretizing multi-dimensional units from ecological gradients that are inherently continuous (Halvorsen et al 2020). This challenge is common to all pattern resolution (Levin 1992) but arguably amplified in ecosystem ecology. We do not propose a single analytical solution to this challenge, for it is embodies aspects of theory, philosophy, and pragmatism.

Perspective

We frame ESPM as an outgrowth in the advancement of spatial ecosystem ecology. As this latter branch of ecology continues to evolve, efforts to model within- and cross-ecosystem patterns and processes are increasingly being tested in complex real-world study landscapes. This is where the potential to bridge theoretical and applied modelling traditions is greatest. While most spatial ecosystem models, at this frontier, are formulated to predict responses to ecosystem heterogeneity, we focus on the origins and organization of that heterogeneity. Our framework is intended to help encourage new strategic directions in the type of integrative modelling needed to strengthen conservation assessments of ecosystems, at those region-specific extents relevant to environmental decision making (Perino et al 2022). While forecasting ecosystem change in future scenarios is essential for decision making, quantifying current ecosystem patterns is comparably important (Watson et al 2020). Outputs of our framework can also help catalyze efforts needed to improve basic understanding of structural, compositional, and functional relationships across scales and among extant ecosystems.

Our approach is not intended as an endpoint. Instead, we offer an alternative analytical perspective to quantify emergent aspects of ecosystems, and how they assemble across space. Patterns of heterogeneity reflect past and present interactions among biotic and abiotic elements. The unique combination of ecosystem elements and processes, expressed at a specific time and place, is like a dynamic blueprint. We do not model dynamic aspects of that blueprint, or assembly processes themselves. Our focus is on the spatially structured determinants and outcomes established by those processes. We do not view this apparent divergence from the functional or mechanistic (often referred to as process-based) emphasis typical in ecosystem and meta-ecosystem ecology, as a disciplinary incongruency. Pattern and process span all levels of organization in nature, including ecosystems. Numerous spatial models have been formulated to predict and help understand biodiversity distributional patterns (Araujo et al 2019), but spatial models of ecosystem pattern are scarce (Geary et al 2020) and frequently absent from broad-scoped thematic reviews. Even in extensive reviews (e.g., Zurell et al 2021) of spatial biodiversity models, only process-based spatial ecosystem models are sufficiently common for consideration. These trends persist in general ecology almost twenty years after Loreau et al (2003) highlighted the schism between models of biotic (e.g., population, community) elements, and those built for ecosystem prediction. Our approach is intended to help provide a common basis for linking predictions across scales and levels of ecological organization, including constituent facets of biotic and abiotic complexity.

Recognition of the importance of abiotic pattern and process for understanding ecosystems, and other components of biodiversity, has largely originated within the geoscience community (Richter and Billings 2015). Here, biogeochemical elements are often emphasized, while other physical and more strictly abiotic variables and processes are rarely addressed as completely (Johnson and Martin 2016, Dor-Haim et al 2019). This disparity in purview has affected perceptions of what an ecosystem is, how it can be modelled, and which abiotic variables should be considered in empirical ecosystem research (Richter and Billings 2015, Johnson and Martin 2016). We suggest abiotic complexity is not singularly important because biota rely on it – it is fundamental to ecosystem organization. Moreover, many abiotic variables and properties originate, or are shaped by, from biotic processes (e.g., humus, water chemistry, soil structure). Recognition of abiotic components of ecosystem constitution and dynamics, and their influence on ecosystem properties has begun to rise within the ecological research community (van der Plas 2019). In addition, recent perspectives and methodological guides (e.g., Zarnetske et al 2019, Alahuhta et al 2020, Hjort et al 2022) hold considerable promise for increasing broader awareness of abiotic complexity, and for catalyzing new approaches for understanding joint biotic-abiotic patterns and processes shaping ecosystems.

In what follows, we outline recommendations for moving forward. While our case study includes the kind of detailed workflow deemed essential to advance predictive ecology (McIntire et al 2022), much effort remains to explore the strengths, and to identify and address the limitations, of ESPM as an analytical template for predicting ecosystem organization. We summarize three overarching issues required to refine our framework; model data and testing, linking model and theory refinement, and applying model outputs to support environmental decision making.

Data required for spatial modelling of species and communities are generally well documented (Bryn et al 2021). The extraordinary popularity of these models has also prompted development of new training and predictor data, and investigations of the suitability these data have for biodiversity modelling are ongoing (e.g., Guillera-Arroita et al 2015, Brodie et al 2020). However, comparable overviews of basic data necessary for ESPM, and similar spatial ecosystem models, are rare and standardized guidelines for data selection do not exist. Prominent advances in spatial ecosystem modelling largely center on the application of remote sensing data for predictions of individual ecosystem components or properties, such as structure (e.g., topography, vegetation physiognomy - D'Urban Jackson et al 2020) or function (e.g., productivity, disturbance – Anderson 2018). For the most part, these latter spatial models do not incorporate ground-derived field measures of biotic and abiotic ecosystem compositional and structural complexity, although recent efforts promote better integration across varied data sources (e.g., Pasetto et al 2018,). Instead, such field data are typically applied for independent spatial predictions of population, community, or geophysical components of nature. Data prerequisite for ESPM include both the detailed type of in situ ecosystem data collected as part of ecosystem inventories, or long-term monitoring, coupled with specific predictors, many of them remotely derived, necessary for modelling ecosystem gradients. Environmental predictors employed for spatial modelling of other components of biodiversity and geodiversity can be adapted for this latter purpose (Simensen et al 2020).

To facilitate overall improvement of our framework, we suggest adoption of standardized data compilation and processing protocols, comparable to those recently established for spatial biodiversity models (Zurell et al 2020). Test and predictor data should be made available directly or via data extraction pipelines provided in study documentation. These provisions will not only help ensure model repeatability and interoperability but facilitate efforts to transfer models to new locations or times. These measures will be particularly helpful where data are sparse or absent, or where environmental circumstances are significantly dissimilar. Furthermore, records of data, metadata, and data processing routines are necessary to determine the influence spatial context has on predictive outcomes.

To make models useful, they must be reliable, flexible, and informative. Many models are only applicable in specific circumstances, compromising generality or predictive inference for accuracy. Predictive models are usually improved by testing with different data, parameters, performance metrics, and algorithms. And ultimately, the only way to improve predictive capacity, and to refine modelling approaches, is to predict frequently and to implement model assessment protocols (Dietze 2017, McIntire et al 2022). We suggest ESPM models need to be tested in different biomes, at different landscape extents, and with different subsets of total regional ecosystem variation. As realistic ecosystem models tend to be more complex (Evans 2012, Mouquet et al 2015), often creating computational challenges, reduced datasets from smaller, less complex, regions are recommended for initial model testing. Extending predictions to unsurveyed areas, as part of testing routines, may be informed by new methods intended to identify the area of applicability for predictive spatial models (Meyer and Pebesma 2021). Lastly, running modelling trials with simulated data can help simplify and expediate testing scenarios before applying them to real data. Simulated data can also be employed to help evaluate model performance, as part of assessment trials (Gallagher et al 2021).

Theories established to explain ecosystem organization have been closely tied with efforts to quantify pattern within and among ecosystems (Levin 1992). And while predictions typically accompany theory development, they also serve to corroborate or refine theory (Houlahan et al 2017). However, ecological theories are often incomplete and fragmented (Mouquet et al 2015) and models may employ elements of multiple theories with different origins. Our predictive framework draws from several interrelated theories all coupling aspects of both static and dynamic pattern resolution. Process-based approaches for predicting ecosystem dynamics are essential, but they can be greatly strengthened by efforts to resolve contemporary ecosystem organizational complexities, many of which remain vague and inadequately understood. Moreover, process-based (mechanistic) and static pattern-based (phenomenological or correlative) models lie along a continuum, sharing many of the same theoretical underpinnings and, in some cases, overlapping research objectives (Mouquet et al 2015). We suggest any model along this continuum can serve as a starting point to abstract complexity, further understanding, and test theory. Such models are different representations of the same ecological entity, and we posit greater effort is required to reconcile and facilitate knowledge exchange among various ecosystem modelling strategies. To help advance this synthesis, we further suggest ESPM outputs could be applied as inputs in process-based (e.g., simulated) spatial ecosystem models. This approach could help determine how spatial ecosystem patterns respond under dynamic scenarios.

We close with a note on matching models to purpose. Part of our motivation with this Forum has been to promote a novel predictive approach for guiding ecosystem-based environmental decision making. Regulatory authorities and conservation practitioners need reliable, timely, and accessible evidence to inform and evaluate their actions. Many decisions are made on short time horizons, tied to fluctuating budgetary and political circumstances. To underpin these decisions, tactical ecosystem predictions are required to determine which individual ecosystems occur where and why; how ecosystems relate to one another (i.e., spatially, functionally, biotically, and abiotically) within their respective ranges; what environmental drivers are important across varied spatial and temporal contexts; which ecosystems support priority constituents (e.g., rare species, functional traits, carbon) or properties (e.g., productivity); and which ecosystems are declining and or otherwise vulnerable. These aspects of contemporary ecosystem geography reflect past assembly pathways, including the consequences of natural disturbances and human land use, and provide a basis to improve strategic predictions of ecosystem change. Tools which facilitate synthesis of ecosystem organization and dynamics are critical for successful natural resource management because decision makers must be mindful of both existing and projected circumstances when settling trade-offs among conflicting land uses and societal pressures.

References

Alahuhta, J., Toivanen, M. and Hjort, J., 2020. Geodiversity-biodiversity relationship needs more empirical evidence. Nature Ecology and Evolution 4, pp.2–3.

Anderson, C.B., 2018. Biodiversity monitoring, earth observations and the ecology of scale. Ecology letters, 21(10), pp.1572-1585.

Anderson, S.C., Elsen, P.R., Hughes, B.B., Tonietto, R.K., Bletz, M.C., Gill, D.A., Holgerson, M.A., Kuebbing, S.E., McDonough MacKenzie, C., Meek, M.H. and Verissimo, D., 2021. Trends in ecology and conservation over eight decades. Frontiers in Ecology and the Environment; doi: 10.1002/fee.2320

Araujo, M.B., Anderson, R.P., Marcia Barbosa, A., Beale, C.M., Dormann, C.F., Early, R., Garcia, R.A., Guisan, A., Maiorano, L., Naimi, B. and O'Hara, R.B., 2019. Standards for distribution models in biodiversity assessments. Sci. Adv. 5, pp.1–12.

Austin, M.P., 2013. Vegetation and environment: discontinuities and continuities. In: van der Maarel, E. and Franklin, Vegetation Ecology, Second Edition. John Wiley and Sons, Ltd.

Baveye, P.C., Chalhoub, M., Choquet, P., and Montagne, D., 2018. Is the focus on "ecosystems" a liability in the research on nature's services? Front. Ecol. Evol. 6:226. doi: 10.3389/fevo.2018.00226

Blew, R.D., 1996. On the definition of ecosystem. Ecological Society of America Bulletin 77, pp.171–173.

Bonan, G.B. and Doney, S.C., 2018. Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models. Science 359, eaam8328.

Brodie, S.J., Thorson, J.T., Carroll, G., Hazen, E.L., Bograd, S., Haltuch, M.A., Holsman, K.K., Kotwicki, S., Samhouri, J.F., Willis-Norton, E and Selden, R.L., 2020. Trade-offs in covariate selection for species distribution models: a methodological comparison. Ecography, 43(1), pp.11-24.

Bryn, A., Bekkby, T., Rinde, E., Gundersen, H. and Halvorsen, R., 2021. Reliability in distribution modeling—A synthesis and step-by-step guidelines for improved practice. Frontiers in Ecology and Evolution, pp.830.

Bullock, J.M., Fuentes-Montemayor, E., McCarthy, B., Park, K., Hails, R.S., Woodcock, B.A., Watts, K., Corstanje, R. and Harris, J., 2022. Future restoration should enhance ecological complexity and emergent properties at multiple scales. Ecography, doi: 10.1111/ecog.05780.

Cadenasso, M.L., Pickett, S.T., Weathers, K.C. and Jones, C.G., 2003. A framework for a theory of ecological boundaries. BioScience, 53(8), pp.750-758.

Caron-Lormier, G., Bohan, D.A., Hawes, C., Raybould, A., Haughton, A.J., Humphry, R.W., 2009. How might we model an ecosystem? Ecol. Model. 220, pp.1935–1949.

Convention on Biological Diversity, 2021. First draft of the post-2020 global biodiversity framework. United Nations Convention on Biological Diversity.

Chao, A. and Jost, L., 2012. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. Ecology, 93(12), pp.2533-2547.

Cogbill, C.V., Thurman, A.L., Williams, J.W., Zhu, J., Mladenoff, D.J. and Goring, S.J., 2018. A retrospective on the accuracy and precision of plotless forest density estimators in ecological studies. Ecosphere, 9(4), p.e02187.

Currie, W.S., 2011. Units of nature or processes across scales? The ecosystem concept at age 75. New Phytol. 190, pp.21 –34.

D'Amen, M.D., Rahbek, C., Zimmerman, N.E., Guisan, A. 2017. Spatial predictions at the community level: from current approaches to future frameworks. Biol. Rev., 92, pp.169 − 187.

Dansereau, G., Legendre, P. and Poisot, T., 2022. Evaluating ecological uniqueness over broad spatial extents using species distribution modelling. Oikos, 2022(5), p.e09063.

DeAngelis, D. L., and Yurek, S. 2016. Spatially explicit modeling in ecology: A review. Ecosystems, 20, pp.284–300.

Dietze, M., 2017. Ecological Forecasting. Princeton Univ Press, Princeton.

Di Minin, E., Correia, R.A. and Toivonen, T., 2021. Quantitative conservation geography. Trends in Ecology & Evolution 37, pp.42–52.

Dor-Haim, S., Orenstein, D.E. and Shachak, M., 2019. Web of interactions among diversity approaches to identify ecosystem essential variables: Negev Highlands case study. Ecosphere 10(11):e02906. 10.1002/ecs2.2906

D'Urban Jackson, T., Williams, G.J., Walker-Springett, G. and Davies, A.J., 2020. Three-dimensional digital mapping of ecosystems: a new era in spatial ecology. Proceedings of the Royal Society B, 287(1920), p.20192383.

Evans, M. R., 2012. Modelling ecological systems in a changing world. Phil. Trans. R. Soc. B 367, pp181–190.

Ferrier, S. and Guisan, A., 2006. Spatial modelling of biodiversity at the community level. J. Appl. Ecol. 43, pp.393–404.

Frazier, A.E., 2022. Scope and its role in advancing a science of scaling in landscape ecology. Landscape Ecology, pp.1-7.

Fulton, E.A., Blanchard, J.L., Melbourne-Thomas, J., Plaganyi, E.E. and Tulloch, V.J., 2019. Where the Ecological Gaps Remain, a Modelers' Perspective. Front. Ecol. Evol. 7, pp.44–56

Gallagher, C.A., Chudzinska, M., Larsen-Gray, A., Pollock, C.J., Sells, S.N., White, P.J. and Berger, U., 2021. From theory to practice in pattern-oriented modelling: identifying and using empirical patterns in predictive models. Biological Reviews, 96(5), pp.1868-1888.

Geary, W.L., Bode, M., Doherty, T.S., Fulton, E.A., Nimmo, D.G., Tulloch, A.I., Tulloch, V.J. and Ritchie, E.G., 2020. A guide to ecosystem models and their environmental applications. Nat. Ecol. Evol. 4, pp.1459–1471

Gignoux, J., Davies, I., Flint, S., Zucker, J-D., 2011. The ecosystem in practice: Interest and problems of an old definition for constructing ecological models. Ecosystems 14(7), pp.1039–1054.

Gounand, I., Harvey, E., Little, C.J. and Altermatt, F., 2018. Meta-ecosystems 2.0: rooting the theory into the field. Trends in Ecology & Evolution, 33(1), pp.36-46. Guisan, A., Thuiller, W., Zimmermann, N.E., 2017. Habitat Suitability and Distribution Models: With Applications in R. Cambridge University Press.

Guillera-Arroita, G., Lahoz-Monfort, J.J., Elith, J., Gordon, A., Kujala, H., Lentini, P.E., McCarthy, M.A., Tingley, R. and Wintle, B.A., 2015. Is my species distribution model fit for purpose? Matching data and models to applications. Global Ecology and Biogeography, 24(3), pp.276-292.

Guisan, A., Thuiller, W., Zimmermann, N.E. 2017. Habitat Suitability and Distribution Models: With Applications in R. Cambridge University Press.

Halvorsen, R., Skarpaas, O., Bryn, A., Bratli, H., Erikstad, L., Simensen, T. and Eriksen, E.L., 2020. Towards a systematics of ecodiversity: the EcoSyst framework. Global Ecol. Biogeogr. 10.1111/geb.13164

Harris, D. J., 2015. Generating realistic assemblages with a joint species distribution model. Methods in Ecology and Evolution 6, pp.465–473.

Harvey, E., Marleau, J.N., Gounand, I., Leroux, S.J., Firkowski, C.R., Altermatt, F., Blanchet, F.G., Cazelles, K., Chu, C., d'Aloia, C. and Donelle, L., 2021. A general meta-ecosystem model to predict ecosystem function at landscape extents. hal-03407501

Hill, N., Woolley, S.N., Foster, S., Dunstan, P.K., McKinlay, J., Ovaskainen, O. and Johnson, C., 2020. Determining marine bioregions: A comparison of quantitative approaches. *Methods in Ecology and Evolution* , *11* (10), pp.1258-1272.

Hjort, J., Tukiainen, H., Salminen, H., Kemppinen, J., Kiilunen, P., Snare, H., Alahuhta, J. and Maliniemi, T., 2022. A methodological guide to observe local-scale geodiversity for biodiversity research and management. Journal of Applied Ecology. 00: pp.1-13.

Holling, C.S., 1992. Cross-scale morphology, geometry and dynamics of ecosystems. Ecological Monographs 62, 447–502.

Houlahan, J., McKinney S., Anderson T., McGill B. 2017. The priority of prediction in ecological understanding. Oikos 126: pp.1–7.

Hui, F.K.C., Taskinen, S., Pledger, S., Foster, S.D. and Warton, D.I., 2015. Model-based approaches to unconstrained ordination. Methods in Ecology and Evolution, 6, pp.399–411.

Johnson, E.A. and Martin, Y.E. (eds.), 2016. A Biogeoscience Approach to Ecosystems, Cambridge Univ. Press, UK, 2016.

Jorgensen, S.E., 2016. Ecological Model Types. Developments in Environmental Modelling, 28. Elsevier

Keith, D.A., Rodriguez, J.P., Rodriguez-Clark, K.M., Nicholson, E., Aapala, K., Alonso, A., Asmussen, M., Bachman, S., Basset, A., Barrow, E.G. and Benson, J.S., 2013. Scientific Foundations for an IUCN Red List of Ecosystems. PLoS ONE 8(5): e62111. doi:10.1371/ journal.pone. 0062111

Keith, D.A., Ferrer-Paris, J.R., Nicholson, E., Bishop, M.J., Polidoro, B.A., Ramirez-Llodra, E., Tozer, M.G., Nel, J.L., Mac Nally, R., Gregr, E.J. and Watermeyer, K.E., 2022. A function-based typology for Earth's ecosystems. Nature, pp.1-6.

Kissling, W.D., Walls, R., Bowser, A., Jones, M.O., Kattge, J., Agosti, D., Amengual, J., Basset, A., Van Bodegom, P.M., Cornelissen, J.H. and Denny, E.G., 2018. Towards global data products of Essential Biodiversity Variables on species traits. Nature ecology and evolution, 2(10), pp.1531-1540.

Konig, C., Weigelt, P., Schrader, J., Taylor, A., Kattge, J. and Kreft, H., 2019. Biodiversity data integration—the significance of data resolution and domain. PLoS biology, 17(3), p.e3000183.

Konig, C., Wuest, R.O., Graham, C.H., Karger, D.N., Sattler, T., Zimmermann, N.E. and Zurell, D., 2021. Scale dependency of joint species distribution models challenges interpretation of biotic interactions. Journal of Biogeography, 48(7), pp.1541-1551.

Levin, S.A., 1992. The problem of pattern and scale in ecology. Ecology, 73(6), pp.1943–1967.

Leroux, S.J., and Loreau, M., 2012. Dynamics of reciprocal pulsed subsidies in local and meta-ecosystems. Ecosystems 15, pp.48-59.

Leroux, S.J., Wal, E.V., Wiersma, Y.F., Charron, L., Ebel, J.D., Ellis, N.M., Hart, C., Kissler, E., Saunders, P.W., Moudra, L. and Tanner, A.L., 2017. Stoichiometric distribution models: Ecological stoichiometry at the landscape extent. Ecol. Lett. 20, 1495–1506. doi: 10.1111/ele.12859

Loreau, M., 2010. Linking biodiversity and ecosystems: towards a unifying ecological theory. Philos. Trans. R. Soc. B: Biol. Sci.365, pp.49–60.

Loreau, M., 2020. Ecosystem. In: Dobson, A. Tilman, D. and Holt, R. D. (Eds.), Unsolved Problems in Ecology. Princeton University Press, Princeton, USA.

Loreau, M., Mouquet, N. and Holt, R.D., 2003. Meta-ecosystems: a theoretical framework for a spatial ecosystem ecology. Ecol. Lett., 6, pp.673–679.

Lyons, M.B., Foster, S.D. and Keith, D.A., 2017. Simultaneous vegetation classification and mapping at large spatial scales. Journal of Biogeography, 44(12), pp.2891-2902.

Massol, F., Gravel, D., Mouquet, N., Cadotte, M., Fukami, T. and Leibold, M., 2011. Linking community and ecosystem dynamics through spatial ecology. Ecology Letters,14, pp.313–323.

McIntire, E.J., Chubaty, A.M., Cumming, S.G., Andison, D., Barros, C., Boisvenue, C., Hache, S., Luo, Y., Micheletti, T. and Stewart, F.E., 2022. PERFICT: A Re-imagined foundation for predictive ecology. Ecology Letters.

Metzger, M.J., Brus, D.J., Bunce, R.G.H., Carey, P.D., Goncalves, J., Honrado, J.P., Jongman, R.H.G., Trabucco, A. and Zomer, R., 2013. Environmental stratifications as the basis for national, European and global ecological monitoring. Ecological Indicators, 33, pp.26-35.

Meyer, H. and Pebesma, E., 2021. Predicting into unknown space? Estimating the area of applicability of spatial prediction models. Methods in Ecology and Evolution, 12(9), pp.1620-1633.

Miller, J.R., Turner, M.G., Smithwick, E.A., Dent, C.L. and Stanley, E.H., 2004. Spatial extrapolation: the science of predicting ecological patterns and processes. BioScience, 54(4), pp.310-320.

Mod, H.K., Chevalier, M., Luoto, M. and Guisan, A., 2020. Scale dependence of ecological assembly rules: Insights from empirical datasets and joint species distribution modelling. Journal of Ecology, 108(5), pp.1967-1977.

Mokany, K., Ferrier, S., Connolly, S.R., Dunstan, P.K., Fulton, E.A., Harfoot, M.B., Harwood, T.D., Richardson, A.J., Roxburgh, S.H., Scharlemann, J.P. and Tittensor, D.P., 2016. Integrating modelling of biodiversity composition and ecosystem function. Oikos 125, pp.10–19

Mokany, K., Ware, C., Woolley, S.N., Ferrier, S. and Fitzpatrick, M.C., 2022. A working guide to harnessing generalized dissimilarity modelling for biodiversity analysis and conservation assessment. Global Ecology and Biogeography 00: pp.1-20.

Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputie, A., Eveillard, D., Faure, D., Garnier, E., Gimenez, O., Huneman, P. and Jabot, F., 2015. Predictive ecology in a changing world. Journal of Applied Ecology 52: pp.1293-1310.

Nash, K.L., Allen, C.R., Angeler, D.G., Barichievy, C., Eason, T., Garmestani, A.S., Graham, N.A., Granholm, D., Knutson, M., Nelson, R.J. and Nystrom, M., 2014. Discontinuities, cross-scale patterns, and the organization of ecosystems. Ecology, 95, pp.654–667.

Newman, E.A., Kennedy, M.C., Falk, D.A. and McKenzie, D., 2019. Scaling and complexity in landscape ecology. Frontiers in Ecology and Evolution, 7, pp.293.

Nieto-Lugilde, D., Maguire, K. C., Blois, J. L., Williams, J. W., and Fitzpatrick, M. C. 2018. Multiresponse algorithms for community-level modelling: Review of theory, applications, and comparison to species distribution models. Methods in Ecology and Evolution, 9, pp.834 – 848.

Norberg, A., Abrego, N., Blanchet, F. G., Adler, F. R., Anderson, B. J., Anttila, J., Ovaskainen, O., 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. Ecological Monographs, 89(3), e01370. https://doi.org/10.1002/ecm.1370

O'Neill, R.V. 2001. Is it time to bury the ecosystem concept? (with full military honors, of course!). Ecology 82, pp.3275–84.

Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F.G., Duan, L., Dunson, D., Roslin, T., Abrego, N., 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. Ecol. Lett. 20, pp.561–576. (doi:10. 1111/ele.12757)

Ovaskainen, O. and Abrego, N., 2020. Joint Species Distribution Modelling – With Applications in R. Cambridge University Press.

Pasetto, D., Arenas-Castro, S., Bustamante, J., Casagrandi, R., Chrysoulakis, N., Cord, A.F., Dittrich, A., Domingo-Marimon, C., El Serafy, G., Karnieli, A. and Kordelas, G.A., 2018. Integration of satellite remote sensing data in ecosystem modelling at local scales: Practices and trends. Methods in Ecology and Evolution, 9(8), pp.1810-1821.

Perino, A., Pereira, H.M., Felipe-Lucia, M., Kim, H., Kuhl, H.S., Marselle, M.R., Meya, J.N., Meyer, C., Navarro, L.M., van Klink, R. and Albert, G., 2022. Biodiversity post-2020: Closing the gap between global targets and national-level implementation. Conservation Letters. 2022; e12848. https://doi.org/10.1111/conl.12848

Peterson, G.D. 2002. Contagious disturbance, ecological memory, and the emergence of landscape pattern. Ecosystems 5, 329–338

Pickett, S.T.A. and Cadenasso, M.L., 1995. Landscape ecology: Spatial heterogeneity in ecological systems. Science 269. pp.331–334.

Pickett, S.T.A. and Cadenasso, M.L., 2002. The ecosystem as a multidimensional concept: meaning, model, and metaphor. Ecosystems 5, pp.1-10.

Pinto-Ledezma, J.N. and Cavender-Bares, J., 2021. Predicting species distributions and community composition using satellite remote sensing predictors. Scientific Reports, 11(1), pp.1-12.

Poggiato, G., Munkemuller, T., Bystrova, D., Arbel, J., Clark, J.S. and Thuiller, W., 2021. On the interpretations of joint modeling in community ecology. Trends in Ecology & Evolution, 36(5), pp.391-401.

Pollock, L.J., O'connor, L.M., Mokany, K., Rosauer, D.F., Talluto, M.V. and Thuiller, W., 2020. Protecting Biodiversity (in all its Complexity): New Models and Methods. Trends in Ecology & Evolution 35(12), pp.1119–1128

Regos, A., Gagne, L., Alcaraz-Segura, D., Honrado, J.P. and Dominguez, J., 2019. Effects of species traits and environmental predictors on performance and transferability of ecological niche models. Scientific reports, 9(1), pp.1-14.

Richter, D.B., and Billings, S.A., 2015. "One physical system": Tansley's ecosystem as Earth's critical zone. New Phytol 206, pp.900–912. https://doi.org/10.1111/nph.13338.

Rowe, J.S., 1961. The level of integration concept and ecology. Ecology 42, pp.420-427.

Rowe, J.S., and Barnes, B. V., 1994. Geo-ecosystems and bio-ecosystems. Bulletin of the Ecological Society of America, Vol. 75, No. 1, pp. 40-41.

Sayre, R., Comer, P, Warner H, Cress, J. 2009. A new map of standardized terrestrial ecosystems of the conterminous United States. US Geological Survey Professional Paper 1768

Shea, T., Woodley, A., Lindsay, K., Lounds, J., Worbets, B., St. Hilaire, A., Sumner, J., 2018. Canada's conservation vision: A report of the national advisory panel. Ottawa, ON: Government of Canada.

Simensen, T., Horvath, P., Vollering, J., Erikstad, L., Halvorsen, R., and Bryn, A., 2020. Composite landscape predictors improve distribution models of ecosystem types. Diversity and Distributions. 26(8), pp.928–943. doi: 10.1111/ddi.13060

Sonnewald, M., Dutkiewicz, S., Hill, C. and Forget, G., 2020. Elucidating ecological complexity: Unsupervised learning determines global marine eco-provinces. Science Advances, 6(22), p.eaay4740.

Soranno, P.A., Wagner, T., Collins, S.M., Lapierre, J.F., Lottig, N.R. and Oliver, S.K., 2019. Spatial and temporal variation of ecosystem properties at macroscales. Ecology Letters, 22(10), pp.1587-1598. Svirezhev, Y.M., Krysanova, V.P., Voinov, A.A., 1984. Mathematical modelling of a fishpond ecosystem. Ecol. Model. 21, pp.315 – 337.

Svirezhev, Y.M., Krysanova, V.P. and Voinov, A.A., 1984. Mathematical modelling of a fish pond ecosystem. *Ecological modelling* ,*21* (4), pp.315-337.

Taranu, Z.E., Pinel-Alloul, B. and Legendre, P., 2021. Large-scale multi-trophic co-response models and environmental control of pelagic food webs in Quebec lakes. Oikos, 130(3), pp.377-395.

Tikhonov, G., Duan, L., Abrego, N., Newell, G., White, M., Dunson, D. and Ovaskainen, O., 2020. Computationally efficient joint species distribution modelling of big spatial data. Ecology, e02929. https://doi.org/10.1002/ecy.2929

van der Plas, F., 2019. Biodiversity and ecosystem functioning in naturally assembled communities. Biological Reviews, 94(4), pp.1220-1245.

Warton, D.I., Blanchet, F.G., O'Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C. and Hui, F.K., 2015. So many variables: joint modeling in community ecology. Trends in Ecology & Evolution, 30(12), pp.766-779.

Watson, J.E., Keith, D.A., Strassburg, B.B., Venter, O., Williams, B. and Nicholson, E., 2020. Set a global target for ecosystems. Nature 578, pp.360–362.

Wilkinson, D. P., Golding, N., Guillera-Arroita, G., Tingley, R. and McCarthy, M.A., 2021. Defining and evaluating predictions of joint species distribution models. Methods Ecol. Evol. 12, pp.394–404.

Xia, J., McGuire, A.D., Lawrence, D., Burke, E., Chen, G., Chen, X., Delire, C., Koven, C., MacDougall, A., Peng, S. and Rinke, A., 2017. Terrestrial ecosystem model performance in simulating productivity and its vulnerability to climate change in the northern permafrost region. Journal of Geophysical Research: Biogeosciences, 122(2), pp.430-446.

Yang, Q., F.-R. Meng, C. P.-A. Bourque and Z. Zhao., 2017. Production of high-resolution forest-ecosite maps based on model predictions of soil moisture and nutrient regimes over a large forested area. Sci. Rep. 7(1), pp1–13. doi:10.1038/s41598-017-11381-z.

Zarnetske, P.L., Read, Q.D., Record, S., Gaddis, K.D., Pau, S., Hobi, M.L., Malone, S.L., Costanza, J., M. Dahlin, K., Latimer, A.M. and Wilson, A.M., 2019. Towards connecting biodiversity and geodiversity across scales with satellite remote sensing. Global Ecology and Biogeography, 28(5), pp.548-556.

Zhang, C., Chen, Y., Xu, B., Xue, Y. and Ren, Y., 2018. Comparing the prediction of joint species distribution models with respect to characteristics of sampling data. Ecography, 41(11), pp.1876-1887.

Zhao, Q., Basher, Z. and Costello, M.J., 2020. Mapping near surface global marine ecosystems through cluster analysis of environmental data. Ecological Research, 35(2), pp.327-342.

Zurell, D., Franklin, J., Konig, C., Bouchet, P.J., Dormann, C.F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A. and Lahoz-Monfort, J.J., 2020. A standard protocol for reporting species distribution models. Ecography, 43(9), pp.1261-1277.

Zurell, D., Konig, C., Malchow, A.K., Kapitza, S., Bocedi, G., Travis, J. and Fandos, G., 2022. Spatially explicit models for decision-making in animal conservation and restoration. Ecography, 2022(4). doi: 10.1111/ecog.05787.

Supporting Materials

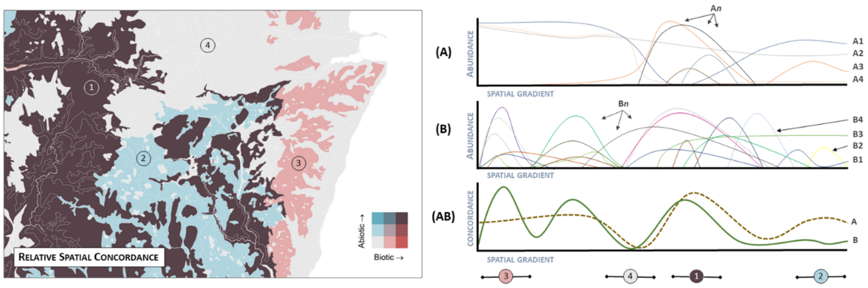
Box 1. Spatial concordance – a basis for resolving intra- and inter-ecosystem patterns

Ecosystems can be distinguished by predicting their component features (Table S1) and the spatially structured relationships between them. These complex relationships span biotic and abiotic domains, scales, and levels of ecological organization. To resolve these complexities, and how they vary across landscapes, we focus on variations in spatial concordance within and among assemblages of biotic (B) and abiotic (A) variables.

Spatial concordance is a condition where two or more ecosystem features show similar scale-specific responses (e.g., presence, abundance) across shared segments of a geographic gradient (plots A and B – right panel). Concordant spatial relationships can arise between individual biotic (e.g., local species populations – O'Neill 2001) and abiotic (e.g., inorganic elements – Austin 2013) ecosystem constituents. Furthermore, concordances can occur across structural (e.g., vegetation physiognomy – Yang et al 2017), compositional (e.g., community – Pinto-Ledezma and Cavender-Bares 2021), and functional (e.g., productivity – O'Neill 2001) ecosystem features. Most studies have focused on concordance among biotic properties (Soranno et al 2019) and particularly how they relate to community variations (e.g., identity – Austin 2013) or to landscape heterogeneity (Pickett and Cadenasso 1995). Instead, we emphasize spatial concordance within and between biotic and abiotic domains (plot AB – right panel). This integrative approach reflects the reciprocal roles (Gignoux et al. 2011, Richter and Billings 2015, Halvorsen et al 2020) of biotic-abiotic constituents and processes in ecosystem assembly (Keith et al 2022). It also provides a means to parse individual and aggregate (Table S1) sources of ecosystem variation and to identify the strongest co-varying biotic-abiotic relationships.

Variations in spatial concordance reveal relationships among ecosystem features and how they scale up to form different ecosystems and the boundaries separating them. To resolve these relationships and their roles in shaping intra- and inter-ecosystem patterns, we distinguish areas where spatial concordance levels are relatively uniform and continuous, from adjacent areas where these collective properties show pronounced change. The latter correspond with boundaries – some narrow and relatively discrete, others wider and more diffuse (Cadenasso et al 2003) – between neighbouring ecosystems. Distinguishing adjacent ecosystems and their common boundaries requires resolution of joint distinctions in biotic-abiotic composition and structure.

To illustrate these patterns, random point occurrences of biotic and abiotic variables were independently simulated across a categorical landscape. The map (left panel) summarizes how these occurrences relate to one another, showing landscape areas marked by: (1) high biotic and abiotic; (2) moderately high abiotic, low biotic; (3) moderately high biotic, low abiotic; and by; (4) low biotic, low abiotic patterns of spatial concordance. These geographic areas are depicted in plots showing individual (plots A and B – right panel) and shared (plot AB – right panel) biotic and abiotic responses (Y axes) along a spatial gradient (X axis).



Box 2. Traditions in ecosystem conceptualization and their influence on research and modelling

A critical factor in ecosystem research is resolution on what an ecosystem is. Many theoretical and practical examinations of the ecosystem concept have been presented since its inception (e.g., O'Neill et al 2001) and although all have improved use and understanding of the term, disparities among the prevailing philosophic and scientific traditions of ecosystem science persist (Loreau 2020). Blew (1996) outlined three major traditions in ecology, each rooted in relatively distinct perceptions of the ecosystem concept and its primary definitional criteria (see Table S1). Similar traditions persist in contemporary ecosystem science (Baveye et al 2018, Zarnetske et al 2019, Alahuhta et al 2020). Modified after Blew (1996), the three major traditions include:

1) Biotic tradition. The biotic approach to ecosystem science is organism focused and typically emphasizes either species or, more often, communities as the most important expression of ecosystem organization. Abiotic constituents are ranked with lesser importance and may be employed to model or describe biotic patterning. Spatial properties may be explicitly, or implicitly, recognized, and functional relationships are inferred (O'Neill 2001). Such units are sometimes called bio-ecosystems, ecological communities, or community-ecosystems (Rowe and Barnes 1994). A prominent example is the IUCN global ecosystem red list, which emphasizes biotically defined units for its operational definition (Keith et al 2013).

2) Functional tradition. Under the tenets of this approach to ecosystem science, the singular representation of holism and complexity are an ecosystem's functional features (Currie 2011). This approach emphasizes the processes (e.g., respiration, decomposition, material cycling), properties (e.g., energy balance, productivity, trophic structuring), or products (e.g., biomass) resulting from interactions between biotic and abiotic ecosystem components (Gignoux et al 2011). Spatial properties may be explicitly or implicitly recognized (DeAngelis and Yurek 2016), and biotic complexity is often generalized to trophic levels, higher taxonomic ranks, or specific functional groups. Prominent examples of this approach are process-based models of ecosystems at local (e.g., Caron-Lormier *et al.* 2009) and global extents (e.g., Bonan and Doney 2018); ecosystem simulation models (e.g., Xia et al 2017); and mathematical models of individual ecosystems (e.g., Svirezhev et al 1984) or meta-ecosystems (e.g., Leroux and Loreau 2012).

3) Abiotic tradition. Following principles of the abiotic tradition, which is sometimes called the *enduring features*approach, ecosystems are interpreted as ecologically homogenous areas of terrestrial or marine geography (e.g., Zhao et al 2020). While biotic variables may be employed to name, map, or describe ecosystems defined through this tradition, biota are usually viewed as ephemeral products of more enduring abiotic drivers, such as climate, landform, soil, hydrology, and geology. These latter physical or marine geodiversity features are considered the building blocks of ecosystems (Sayre et al 2009). This approach deemphasizes ecosystem function and dynamics, focusing on steady-state expressions of ecological variability. Units developed through the abiotic tradition are sometimes called geo-ecosystems (Rowe and Barnes 1994) and they are most applied at broader spatial scales. Biomes, ecoregions, ecozones, ecosites, and landscape ecosystems are prominent examples of units developed, following this tradition, in the contemporary literature (e.g., Yang et al 2017, Keith et al 2022).

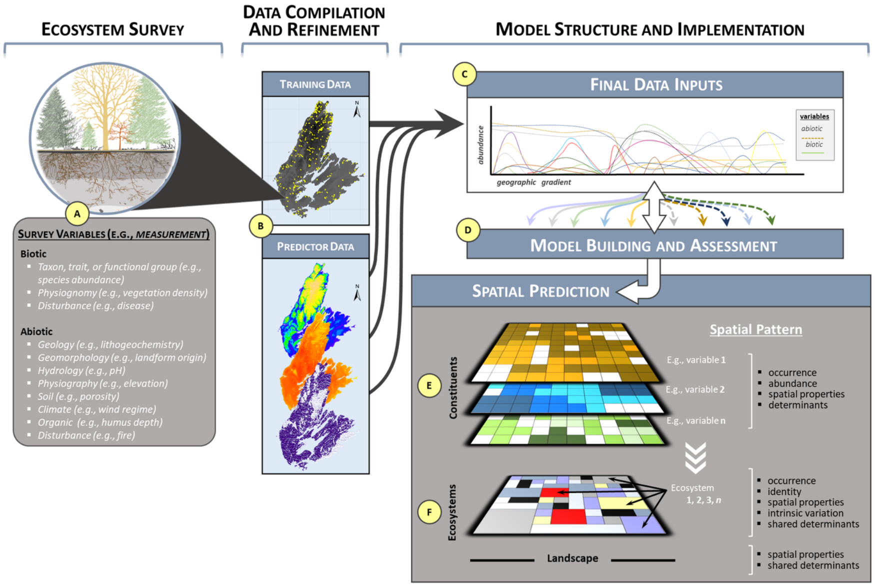


Figure 1. Ecosystem spatial pattern modelling framework components: (A). Ecosystem field surveys, encompassing measures of biotic and abiotic variables. (B). Data Compilation and Refinement. (C) Final model inputs including training and predictor data. (D). Model Building and Assessment. (E). Spatial prediction of ecosystem constituents and properties. (F) Spatial predictions of ecosystem patterns and properties, at local (i.e., within individual ecosystems or ecosystem types) and landscape (i.e., among ecosystems or ecosystem types) extents. E and F adapted from Pollock et al 2020.

Table 1. Practical solutions for common challenges anticipated during ESPM implementation, following workflow components outlined in Figure 1.

Component	**Challenge**	**Solution**
Ecosystem survey (A)	Optimizing survey effort across disparate environmental gradients, scales, and spatial contexts	Implement stratified random survey design (e.g., Metzger et al 2013)
	Fixed spatial grains may bias some survey data	Explore utility of plotless survey methods (Cogbill et al 2018)
	Subjective survey site selection may introduce spatial bias	Coarse-grain (sensu Newman et al 2019) data to common raster grid cells (Miller et al 2004)
Training Data Compilation (B) and Finalization (C)	Measurement scales and units employed for biotic and abiotic field survey variables differ markedly	Convert survey measurements to presence/absence records (König et al 2019)
	Disproportionate ratios of biotic to abiotic variables	Pool species by common ecological traits (Kissling et al 2018)
	Adequacy of survey data	Utilize rarefaction methods (Chao and Jost 2012)

Component	Challenge	Solution
	Effects of spatial survey grain and extent on training data structure	Utilize rarefaction methods (Chao and Jost 2012); test for scale dependencies using independent data sets with differing spatial grains (Mod et al 2020)
	Training data dimensionality impacts model performance	Scale pools of test data across levels of ecological complexity; reduce or eliminate disjunctions (sensu Miller et al 2004) in test data
Predictor Data Compilation (B) and Finalization (C)	Determining relevant conceptual, theoretical, and statistical criteria for selecting predictor data	Draw on key theory from relevant disciples. Select predictors to resolve joint biotic-abiotic patterns and processes shaping ecosystems at various scales
	Mismatched spatial resolution of predictors and response variables (Bryn et al 2021)	Coarse-grain training data to match predictor grain (Newman et al 2019); assess effect of spatial scaling on model performance (König et al 2021)
	Unequal numbers of biotic and abiotic predictors	Employ pilot models to test the effect varying ratios, and combinations, of biotic and abiotic predictors has on ESPM performance (Brodie et al 2020)
	Implications of pooling disparate predictors shaping ecosystem pattern across spatial grains	Determine whether modelled responses are shaped more by local and or regional drivers (e.g., Soranno et al 2019), and whether responses vary among ecosystems and ecosystem constituents
	Implications of grouping predictors of direct, indirect, and resource gradients (Austin 2013)	Clarify expected responses of individual and aggregate ecosystem features (Table S1) to individual environmental gradients (Austin 2013)
	Optimizing combinations of predictors	Explore implications of co-variate trade-off (Brodie et al 2020)
Model Building and Assessment (D)	Model algorithm selection	Seek ESPM algorithms with adequate flexibility, functionality, and predictive capacity. Candidate algorithms may be adapted from community-level (e.g., Nieto-Lugilde et al 2018) or bioregion modelling (e.g., Hill et al 2020) techniques.

Component	Challenge	Solution
	Determining model settings	Settings (e.g., cross-validation, latent variables, random effects, residual associations) depend on selected algorithm and on results of trial model implementations; consult published methodological guidelines (e.g., Ovaskainen and Abrego 2020, Mokany et al 2022)
	Evaluating model fit and performance	Metrics for evaluating model fit and performance vary by algorithm and modelling intent; consult appropriate guidelines (Araújo et al 2019, Zurell et al 2020)
	Translating model prediction to mapped spatial outputs	Continuous (e.g., ordination) or discrete (e.g., region of common profile) post-hoc analytical outputs can be mapped in YUV colour space (e.g., Tikhonov et al 2020)
Intrinsic Predictions (E)	Determine value of lower-order intrinsic (e.g., biotic and abiotic ecosystem constituents and properties) predictions for informing higher-order modelling objectives	Improve understanding of commonalities, potential interactions, and properties of individual variables characterizing ecosystems or sites; disentangle the relative contributions of individual variables to assembly mechanisms; relate model results to existing ecological knowledge
Extrinsic Predictions (F)	Select method for resolving the identities, features, and distribution of disparate ecosystems and ecosystem types	Adapt classification methods and indices (e.g., similarity, clustering, ordination) developed for community-level (Ferrier and Guisan 2006) or bioregion modelling (Hill et al 2020) Explore utility of less common tools: concordance analysis (e.g., Taranu et al 2020), ecological uniqueness indices (e.g., Dansereau et al 2022), or embedding techniques such as T-SNE (e.g., Sonnewald et al 2020).

Table 2. Different algorithms considered for predictive ecosystem modelling.

Modelling algorithm	**Algorithm Notes**
Multivariate regression tree (MRT)	multiresponse extention of classification and regression t
Multiresponse multivariate adaptive regression splines (MMARS)	multiresponse version of multivariate adaptive regressior
Constrained linear ordination (CLO)	similar to generalized linear model (GLM)
Constrained quadratic ordination (CQO)	generalized linear model (GLM) with quadratic respons
Constrained additive ordination (CAO)	similar to generalized additive model (GAM)
Multiresponse artificial neural networks (MANN)	a type of machine learning
Multivariate stochastic neural network (MISTN)	a type of machine learning
Hierarchical Bayesian Modelling (HBM)	a relatively newer application of Bayesian algorithms
Latent variable model (LVM)	Bayesian implementation; overlap with HBM
Generalized dissimilarity modelling (GDM)	use continuous functions to model community patterns
Gradient forest (GF)	community level extension of Random Forest
Region of common profile (RCP)	Less frequently applied in ecology
Generalized joint attribute modelling (GJAM)	not commonly applied; conflicting evidence on predictiv

Table 3. Evaluation of different algorithms considered for predictive ecosystem modelling. Features are drawn from recent reviews (Nieto-Lugilde et al 2018, Norberg et al 2017), methodological guides (Mokany et al 2022), personal communication (Ovaskainen 2021), and from assessments of individual algorithms (e.g., Poggiato et al 2021). Algorithms described in Table 2. A – Multivariate regression tree, B – Multiresponse multivariate adaptive regression splines, C – Constrained linear ordination, D – Constrained quadratic ordination, E – Constrained additive ordination, F – Multiresponse artificial neural networks, G – Multivariate stochastic neural network, H - Hierarchical Bayesian model / Latent variable model, I – Generalized dissimilarity modelling, J – Gradient forest, K – Region of common profile, L – Generalized joint attribute modelling

Features	**Features**
Data Inputs	Abundance records can be employed as training data
	Flexibility to accommodate pools of biotic and abiotic variables as training data
	Categorical and continuous variables can be included as training data
	Categorical and or continuous predictor data can be employed
	Large datasets efficiently handled
	Traits, phylogeny, and other ancillary data can be incorporated
Implementation	Model outputs easy to visualize, understand, and interpret
	Readily available tutorials and examples of code
	Technically challenging to implement and fit models
	Computationally demanding and time consuming to run
	Results can be challenging to interpret
	Inscrutable or “black box” approach
Analytical Properties	Allows for non-linear responses
	Allows for interaction between predictor variables
	Accounts for unexplained co-occurrences
	Uncertainty can be explicitly quantified
	Predicts compositional differences between locations as a continuous function of environmental g
	Dissimilarities between pairs of sites related to differences in environmental conditions and geogr
Performance	Moderate to high predictive power